

Bayesian and Regression Models of Human Information Processing—A Myopic Perspective

WARD EDWARDS

Highway Safety Research Institute, The University of Michigan

When Jim Naylor asked me to write an introduction to the magnificent paper by Paul Slovic and Sarah Lichtenstein that forms the body of this issue of *Organizational Behavior and Human Performance*, I wasn't quite sure whether to thank him or hit him (so I did both). His motive was obvious: the Slovic-Lichtenstein paper isn't quite long enough to fill the issue, but is too long to be combined with an honest paper on a different subject. Moreover, he knew I would accept: their paper is much too good for me to pass up the chance to bask in their reflected glory. But the invitation led me to reflect on the ages of scientists. The new PhD is too busy doing experiments to bother reading broadly; also, he had to do too much of that too recently, and is sick of it. That's Age 1. Age 2 comes when the scientist "matures," which means that he becomes responsible for enough experimental literature to generate contradictions and opponents. These, he has no choice but to read, and in trying to discover why he was right all along he may read quite widely—even desperately. The apex of this stage is reached when he has invested so much time in reading that he must again (for the first time since graduate school) report what he has read, in the process spanking friends and soothing enemies. That is how papers like Slovic and Lichtenstein's are generated. In Age 3, contradictions become less important, mostly because one becomes too busy to generate the facts that generate contradictions. At that stage, one tends to generate points of view, usually miscalled theories. If one is lucky, either at generating ideas or at attracting money, research gets structured around these points of view. Finally, at Age 4, one is too senile for any of these other activities, so one writes histories, introductions, committee reports, and the like. While there is a certain numerical justice in my entering Age 4 at age 44, I'm not at all sure I forgive Jim Naylor for spotting the transition.

Well, I might as well indulge myself in the privileges of age; the most enjoyable of these is that of pointing out the importance of history. (That's enjoyable because the historian can thus once again call attention

to his own long-forgotten Age 1 work, obviously seminal to all that came after.)

Incidentally, I asked Mike O'Connor to read a draft of this introduction; his main complaint was that many of its comments are more likely to be intelligible to someone not working in the field after reading Slovic and Lichtenstein's paper than before reading it. So you can treat this as a branch point: you can go on, or read Slovic and Lichtenstein and then return to here, or read Slovic and Lichtenstein and quit. (The third option seems wisest to me.)

HISTORICAL NOTES

Probably the most underrated psychologist of the 1937–1955 period was Egon Brunswik. It wouldn't surprise me if, 100 years from now, he were seen as the most important psychologist of the first half of the 20th century. (If so, please let me know.) He suffered from a turgid, almost unintelligible writing style, and from ideas about how to do experiments (representative design) that, as I see it, drew an overblown conclusion from completely valid premises. He also suffered from the narrowness of his own substantive research interests. But he saw one thing clearly that seems to have escaped all of his contemporaries except the very few who took the trouble to understand him. He saw that psychology is not only about people who emit behavior—it is also, perhaps more importantly, about the tasks that elicit that behavior. That is, he saw that the task-relevant characteristics of the environment are a necessary part of every process theory in psychology.

That was a profound insight. Its consequences began to be felt only after World War II, and even now many deeply sophisticated psychological theorists fail to recognize how much effort they spend in modelling the task, and how little effort they spend in modelling the behavior of the subject in the task. Some theorists do understand this. Frank Restle, for example, once told me about an occasion in which he had tried some five totally different sets of assumptions about the subject, combined with the same set of assumptions about the nature of the task—and obtained indistinguishable predictions. Mathematical psychologists of the Stanford school often (sadly, not always) separate their axioms into some that describe the task and some that describe the subject. Fame awaits the mathematical psychologist who can figure out a satisfactory way to partition the predictive success of a model between its task-describing components and its man-describing components. My own guess is that most successful models now available are successful exactly because of their success in describing tasks, not people. If so, Brunswik is the father of successful psychological theorizing, since

he was the first to make explicit and clear the fact that modelling tasks is different from modelling people, to hunt for tools for modelling tasks, and to provide linkages between models of tasks and models of people.

Moreover, Brunswik emphasized the fallible, probabilistic nature of the environment, and so of most real-world tasks performed in it. This, it seems to me, was not a great insight itself, but rather the result of the preceding one. No one can look at the world with a task-oriented eye and not see that uncertainty and unpredictability pervade it. That Brunswik saw and emphasized this while other psychological theorists of the time did not is a tribute not to his genius, but rather to the effectiveness of the laboratory blinders others wore—blinders that led them to design experiments and theories that featured completely lawful stimulus-producing environments.

It is a shame that Brunswik was no mathematician (and probably was in at least Age 3) because at or before the time he was hunting for tools with which to model tasks in a fallible, probabilistic environment, and finding only the conventional correlational analyses of then-current statistics, Bruno de Finetti (1937), writing in French in the *Annals de l'Institut Henri Poincaré*, not perhaps the journal most closely scanned by a California-based psychologist of German-speaking origin, was presenting to an uninterested world the basic ideas around which the most sophisticated analysis yet available of uncertainty have been structured. These are, of course, the Bayesian ideas of probability theory. They were anticipated, sort of, a decade earlier by a young philosopher, Frank Plumpton Ramsey (1931), who died, still very young, in a motorcycle accident.

De Finetti's ideas remained virtually unlistened-to even when, during World War II, von Neumann and Morgenstern (1944) needed the idea of maximizing expected utility as an underpinning to their theory of games. But von Neumann and Morgenstern were, for some reason, heard by practically everybody—their book must be the most unread best seller in the social sciences other than the Kinsey report. Economists, political scientists, statisticians, and psychologists all climbed on the decision-theoretic bandwagon and by 1954, emerging from Age 1 to Age 2, I could review an extensive, multidisciplinary literature and thus bring the decision-theoretic ideas to what was by then a fairly attentive psychological world.

Brunswik's students and followers carried on after his death, creating a small, loyal in-group who came close to complete divergence from the main stream of psychology. The focus of their work was at the University of Colorado, where Kenneth Hammond led a devoted and productive group.

The decision theorists were more loquacious, as this introduction more than suggests. Partly because they had strong ties to other, more mathematical disciplines and partly because their natural bent toward application led to consistently easy funding, they produced a great deal of paper and noise. Thanks primarily to my fantastic luck in being tutored by L. J. Savage, psychological decision theory discovered Bayes's theorem and personal probabilities in the early 1960's (Edwards, Lindman, & Savage, 1963). But only in the very late 1960's did decision theorists come to understand what Brunswik had implied in the 1940's and 1950's: that models of tasks are at least as important as models of people. Once decision theorists saw that, they (at least some of them) realized that they had been modelling tasks all their lives without knowing it.

Now Paul Slovic and Sarah Lichtenstein, trained in decision theory in my laboratory but literate enough to learn what I never knew, are trying to unify, or at least marry, the Brunswikian with the decision-theoretic traditions. I'm all for it. It would be fun to be married to the Brunswikians, as well as to the statisticians, operations researchers, economists, industrial engineers, systems analysts, and the like.

I haven't said anything so far about multiple regression or analysis of variance—and Norm Anderson will never forgive me. As straight statistical tools these ideas have a long history, but so far as I know their use in psychological theorizing originates in part within the Brunswikian tradition, and otherwise as a part of the introduction of measurement theory into psychology that happened mostly in the 1960's. That is a story in itself—and not one I know very well. But measurement theory inevitably focuses attention on additive models, and that kind of attention can easily lead to attention to analysis of variance. Norm Anderson has been most active in developing this line of thought.

Most explicit psychological models have, or can be made to take, a form like $R = ax + by + cz + \dots + e$, where R is a response, the abc terms are weighting factors on the stimulus dimensions, the xyz terms are measures of the stimulus dimensions, and e is an error term of some sort. Such models are called additive because terms like $xy + x + y$ don't appear; sometimes, in a weaker sense, because terms like x^2 or e^x don't appear. I find it convenient to use a procrustean bed in which to classify such models. All of them tend to take ordinal properties both of R and of the xyz terms seriously; the question is, which, if any, of the variables have cardinal (interval scale) properties that you should take seriously. If the answer is none of them, you are dealing with a conjoint-measurement-type model. If the answer is all of them, you are dealing with a multiple-regression-type model. (Some Bayesian models

fall in this class.) If the answer is that you believe in cardinal properties of R but only ordinal properties of xyz , you are following the lead of Norman Anderson, who calls what he does functional measurement. (This oversimplifies, since Norm is also interested in transformations on R .) If you believe in cardinal properties of xyz but not of R , you may simply be doing nonparametric statistical inference, or you may be doing some fancy kind of Kruskal-Wallace scaling; for reasons that escape me, the literature on inference, and so the Slovic-Lichtenstein paper, mostly ignores this approach.

DEPENDENT VARIABLES AND NORMATIVE OBJECTIVES

Critics of Tolman and his colleagues of the 1930's accused him of leaving his rat lost in thought in the middle of the maze, equipped with a cognitive map but no rules permitting the rat to translate the information into action. On the whole, that criticism remains valid today, not only of the n th generation followers of Tolman, but also of those who made the criticism, and their followers. Except for that offered by decision theory, the only theory of response available within psychology today, so far as I know, asserts that a rewarded response increases in likelihood of occurrence. Perhaps you can take that idea seriously as an account of how you select among acts; I can't. It can perhaps account for thoughtless, casual actions; it is, for example, a good theory to explain smoking. But it cannot begin to account for significant, reasoned actions, such as giving up smoking. When I play poker, for example, a successful bluff does not increase the probability that I'll bluff again on the next hand. When I write an article, submit it to a journal, and get it accepted, I am very unlikely to write that same article again and submit it again to that same journal. Far too many successful responses are not repeated—often exactly because they were successful. The concepts of stimulus discrimination and response differentiation, elaborated to the point of meaninglessness, are invoked to “explain” your and my endless flexibility in adapting our actions to the detailed circumstances in which they occur. But those concepts, so elaborated, amount to nothing more than saying that the organism does what is appropriate to the circumstances in which he finds himself—Tolman's position. So the descendants of Tolman's critics are in the same maze as Tolman's descendants, lost, not in thought, but in making stimulus discriminations and response differentiations. Is there a difference?

Most contemporary psychologists, dissatisfied with the reinforcement-theory explanation of responding and equally dissatisfied with complete silence on the topic, fall back on a vague kind of functionalism—the organism diagnoses the situation and does what is appropriate. This

retreat has an important virtue: it divides psychology into two problems, one of which (diagnosis) is more or less soluble. But its answer to the theory-of-response problem is exactly Tolman's.

Meanwhile, economists and others couldn't get along without a response theory. What should such a theory consist of? It needs, it seems to me, three parts: a theory about how potential responses are invented for consideration, a theory of the process of consideration by which all responses but one are eliminated, and a theory about how the chosen response is translated into actual muscular activities. About the last, we remain silent, hoping for eventual illumination from the neurophysiologists. The first, the problem of creative thinking, is the most important and least touched problem in psychology. Outside psychology, some ideas that may be relevant to the topic exist (e.g., relevance trees), but they haven't been imported yet, and no one knows how they will work. But the topic of winnowing down the candidate responses from many to one has an elegant formal theory and an elaborate psychological research history, to be found in the *Psychological Abstracts* under headings like Decision Processes or Decision Theory. It really says no more than Tolman did—that the organism will do what is most appropriate to the situation in which he finds himself—but it says it quantitatively, in a way that often permits prior calculation of what action is most appropriate, and just how much more appropriate it is than its nearest competitor.

The main motivation for this response theory (which, in its formal structure, is as old as probability theory itself—at least 300 years old) has been, not theoretical, but practical. The goal has been, not to predict responses, but rather to prescribe them. Such prescriptive or normative theories are indispensable tools for the application of science to everyday affairs; diagnosis is of little value unless the rules leading from it to action are known. The descriptive usefulness of such normative theories is somewhat more than a happy accident, though. While we complain of human irrationality, we would probably also (if not too bigoted) admit that the following question and answer contain a powerful, rather accurate, and often operational theory of human behavior:

Q: What is he doing?

A: He's doing the best he can.

Obviously, the goal of any normative response model is to specify what actions are "the best he can" do. If he then does them, the normative model becomes descriptive.

It is, I think, in what it has to say about responding that decision theory differs most from lens-model and functional-measurement ap-

proaches to diagnosis. Decision theory offers a fully developed theory of response selection, and an explicit and fully quantified description of how diagnoses affect action. (However, like everyone else, decision theorists remain silent about response invention.)

Because of their normative orientation (and their enjoyment of money), decision theorists have tended to be extensively involved with practical problems. So, for that matter, have the correlationists. The practical problems of the correlationists have, however, tended to be diagnostic in nature, in situations in which a choice of a diagnostic category implied a clear course of action. (Is he psychotic or neurotic? Will he graduate from graduate school, if admitted?) When the action implications of a diagnosis are unclear or complicated, the apparatus of decision theory seems to me indispensable.

CORRELATIONS AMONG PREDICTORS

Within the field of diagnosis itself, the Brunswikians were clearly first in worrying about correlations among cues—or, as we call it today, the problems of independence and conditional independence. Unfortunately, the mathematical apparatus of correlation coefficients is inconvenient for the problem—though in the hands of sophisticated and imaginative users, these tools, like any, can be made incredibly effective.

The Bayesians found that problem late. I like to think that they made one important theoretical and one important applied contribution to it. The theoretical contribution is the distinction between independence and conditional independence. Consider two data, D_1 and D_2 , and an exhaustive set of two mutually exclusive hypotheses, H_A and H_B . D_2 is independent of D_1 , as any statistics text will tell you, if and only if $P(D_2) = P(D_2|D_1)$. You will find, if you read further in that statistics text, a great deal of nonsense about the importance of independence of this kind to inference. But it is in fact nonsense. Only by a rare statistical accident can two data, both relevant to the diagnostic question about whether H_A or H_B is true, be independent. In general, if D_1 is relevant, it will change the probability distribution as between H_A and H_B . D_2 , to be relevant, must be more probable on one of those hypotheses than on the other. So that change in the relative probabilities of the hypotheses will change the probability of D_2 , thus violating independence.

Actually, no one wants independence of this kind among data. What you and everyone else want instead is conditional independence. That is, you want it to be true that $P(D_2|H_i) = P(D_2|H_i, D_1)$, for both $i = A$ and $i = B$. In other words, if you knew one hypothesis to be true, then you would want knowledge of D_1 not to affect the probability of D_2 . (The reason why most statistics texts fail to point this out is that when

they write $P(D_2)$ they really mean $P(D_2|H_0)$, where H_0 is the null hypothesis; they are really talking about conditional independence, not independence, all along.)

It is possible to produce data and hypotheses showing all four combinations of conditional and unconditional dependence and independence. Moreover, two data may be conditionally independent for one set of hypotheses and conditionally dependent for a different set; which means that no sampling scheme can by itself ensure (conditional) independence.

All this is not discussed in this form by the correlationists. But they do worry, extensively and sophisticatedly, about the correlations among cues, as well as the correlations between cues and responses. And, by at least 20 years, they got there first.

(My main reason for writing this section is that I know absolutely no place in the published literature where the distinction between conditional and unconditional independence is clearly stated—though Bayesian statisticians must have understood it at least since de Finetti in 1937. Now there's a published reference.)

The Bayesians can claim priority in recognizing one very practical form that the problem of correlations among cues can take: the problem of hierarchical inference systems. Very often the output of one inference system is the input to a higher-order one. The photointerpreter says "I think that blob on that photo is a tank"; the analyst says "According to photo-reconnaissance information, the enemy is moving tanks from X to Y along Route 91"; the higher-level analyst says "Intelligence reveals that the enemy is concentrating tanks at Y ; an attack is therefore imminent"; and so on. It is easy to show that such a system will neglect severe problems resulting from violations of conditional independence; it is intuitively reasonable, and experience indicates that such systems will typically be overconfident at their upper, or output, ends. Methods for coping with the difficulty have been developed, but are not as yet in use (Gettys & Willke, 1969). The same problem of course exists in the environments typical of correlation applications, but I have seen no discussion of it in such contexts.

PROSPECTS FOR FUSION, OR AT LEAST COOPERATION

Slovic and Lichtenstein had a practical objective in writing their paper: to get the Bayesians and the correlationists at least to talk to each other, and perhaps to work together. The first of these objectives is certainly going to occur. Obstacles stand in the way of the second. The kinds of real-world problems natural to Bayesians are less natural to correlationists. The Bayesian approach to the semiautomation of

human inference is highly analytic, requiring fragmented judgments that are later combined by an optimal aggregation rule. The correlationist's most common approach is, instead, to model a skillful judge (as in the bootstrapping technique). Both approaches work; no one has even begun to ask which works better, when.

Much of the more abstract Bayesian research is categorical, lending itself to binomial, or at most multinomial, models. Much correlational research is concerned with continuous cues and often with continuous dependent variables as well. (Of course Bayesian models for such continuous cases exist; they differ only trivially from the traditional correlation models.)

The typical finding of both kinds of research is that men are suboptimal, but the characterization of that suboptimality seems very different in the two traditions, and I know of few proposals for reconciling the differences.

But the differences should be reconciled. In that Slovic and Lichtenstein are right. They've done all they could, given the literature available. In doing so, they have not only served a generation of graduate students preparing for a prelim on cognitive processes; they have also set the rest of us a task. We should not only talk, but talk effectively. We are, after all, working on the same problem, using (I trust) the same organism as subject. If we differ about why men are suboptimal, we should find out what the truth is. If our strategies for coping with that suboptimality are different, we should find out which works better when and then, if possible, work out yet a better strategy, using the best features of each.

Well, that should be enough to fill Jim Naylor's pages and insult my friends. Besides, it says what I believe—what I think Slovic and Lichtenstein would like me to believe, after reading their distinguished synthesis. It only remains for me to wonder why you have bothered to work your way through these pages, instead of proceeding at once to the feast that starts on the next page. Onward.

REFERENCES

- DE FINETTI, B. La prévision: Ses lois logiques, ses sources subjectives. *Annals de l'Institut Henri Poincaré*, 1937, **7**, 1-68. (Translated in H. E. Kyburg, Jr., and H. Smokler (Eds.), *Studies in subjective probability*. New York: Wiley, 1964.)
- EDWARDS, W. The theory of decision making. *Psychological Bulletin*, 1954, **51**, 370-417.
- EDWARDS, E., LINDMAN, H., & SAVAGE, L. J. Bayesian statistical inference for psychological research. *Psychological Review*, 1963, **70**, 193-242.
- GETTYS, C. F., & WILLKE, T. A. The application of Bayes' theorem when the true data state is unknown. *Organizational Behavior and Human Performance*, 1969, **4**, 125-141.

- RAMSEY, F. P. Truth and probability. In F. P. Ramsey (Ed.), *The foundations of mathematics and other logical essays*. New York: Harcourt Brace, 1931. (Reprinted in H. E. Kyburg, Jr., and H. Smokler (Eds.), *Studies in subjective probability*. New York: Wiley, 1964.)
- VON NEUMANN, J., & MORGENSTERN, O. *Theory of games and economic behavior*. Princeton: Princeton Univ. Press, 1944.

RECEIVED: July 27, 1971